

Applicant: Li *et al.*

Due Date: June 14, 2003 (Saturday)

Art Unit: 1642

Examiner: Harris, A.M.

Docket: 0609.4560002

Atty: JAG/KRM/DJN

Exhibit A
09/720,086

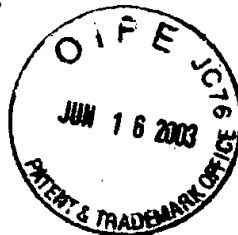
Application No.: 09/720,086

102(e): July 23, 2001

For: *De Novo* DNA Cytosine Methyltransferase Genes, Polypeptides and Uses Thereof

When receipt stamp is placed hereon, the USPTO acknowledges receipt of the following documents:

1. SKGF Cover Letter;
2. Petition For Extension of Time Under 37 C.F.R. § 1.136(a)(1);
3. Fee Transmittal Form;
4. United States Patent & Trademark Office Credit Card Payment Form;
5. Amendment and Reply Under 37 C.F.R. § 1.111;
6. Supplemental Request to Approve Proposed Drawing Corrections with 7 sheets of drawings containing proposed corrections marked in red attached;
7. Letter to PTO Draftsman: Submission of Formal Drawings;
8. Copies of the Formal Drawings (38 sheets); and
9. Return postcard.



Please Date Stamp And Return To Our Courier

Mouse Dnmt3a DNA sequence

1 GAATTCGGG CTGCTGCCG GCGCCCGAC CCGCCGGCC ACACGGCAGA
51 GCGCCTGAA GCCCAGCGCT GAGGCTGCAC TTTCCGAGG GCTTGACATC
101 AGGGTCTATG TTAAAGTCTT AGCTCTTGCT TACAAAGACC ACGGCAATTC
151 CTTCTCTGAA GCCCTCGCAG CCCACAGCG CCCTCGCAGC CCCAGCCTGC
201 CGCCTACTGC CCAGCAATGC CCTCCAGCGG CCGCGGGGAC ACCAGCAGCT
251 CCTCTCTGGA GCGGGAGGAT GATCGAAAGG AAGGAGAGGA ACAGGAGGAG
301 AACCGTGGCA AGGAAGAGCG CCAGGAGCCC AGCGCCACGG CCCGGAAGGT
351 GGGGAGGCCT GGCCGAAGC GCAAGCACCC ACCGGTGGAA AGCAGTGACA
401 CCCCCAAGGA CCCAGCAGTG ACCACCAAGT CTCAGCCCAT GGCCAGGAC
451 TCTGGCCCCT CAGATCTGCT ACCCAATGGA GACTTGAGA AGCGGAGTGA
501 ACCCCAACCT GAGGAGGGGA GCCCAGCTGC AGGCAGAAAG GGTGGGGCCC
551 CAGCTGAAGG AGAGGGAAGT GAGACCCAC CAGAAGCCTC CAGAGCTGTG
601 GAGAATGGCT GCTGTGTGAC CAAGGAAGGC CGTGGAGCCT CTGCAGGAGA
651 GGGCAAAGAA CAGAAGCAGA CCAACATCGA ATCCATGAAA ATGGAGGGCT
701 CCGGGGGCCG ACTGCGAGGT GGCTTGGGCT GGGAGTCCAG CCTCCGTCAG
751 CGACCCATGC CAAGACTCAC CTTCCAGGCA GGGGACCCCT ACTACATCAG
801 CAAACGGAAA CGGGATGAGT GGCTGGCAGG TTGAAAAGG GAGGCTGAGA
851 AGAAAGCCAA GGTAATTGCA GTAATGAATG CTGTGGAAGA GAACCAGGCC
901 TCTGGAGAGT CTCAGAAGGT GGAGGAGGCC AGCCCTCCTG CTGTGCAGCA
951 GCCCACGGAC CCTGCTTCTC CGACTGTGGC CACCACCCCT GAGCCAGTAG
1001 GAGGGGATGC TGGGACAAG AATGCTACCA AAGCAGCCGA CGATGAGCCT
1051 GAGTATGAGG ATGGCCGGGG CTTTGGCATT GGAGAGCTGG TGTGGGGGAA
1101 ACTTCGGGGC TTCTCCTGGT GGCCAGGCCG AATTGTGTCT TGGTGGATGA

FIG. 1A-1

1151 CAGGCCGGAG CCGAGCAGCT GAAGGCACTC GCTGGGTCAT GTGGTTCCGA
1201 GATGGCAAGT TCTCAGTGGT GTGTGTGGAG AAGCTCATGC CGCTGAGCTC
1251 CTTCTGCAGT GCATTCCACC AGGCCACCTA CAACAAGCAG CCCATGTACC
1301 GCAAAGCCAT CTACGAAGTC CTCCAGGTGG CCAGCAGCCG TGCCGGGAAG
1351 CTGTTTCCAG CTTGCCATGA CAGTGATGAA AGTGACAGTG GCAAGGCTGT
1401 GGAAGTGCAG AACAAGCAGA TGATTGAATG GGCCCTCGGT GGCTTCCAGC
1451 CCTCGGGTCC TAAGGGCCTG GAGCCACCAG AAGAAGAGAA GAATCCTTAC
1501 AAGGAAGTTT ACACCGACAT GTGGGTGGAG-CCTGAAGCAG-CTGCTTACGC
1551 CCCACCCCA CCAGCCAAGA AACCCAGAAA GAGCACAACA GAGAAACCTA
1601 AGGTCAAGGA GATCATTGAT GAGCGCACAA GGGAGCGGCT GGTGTATGAG
1651 GTGCGCCAGA AGTGCAGAAA CATCGAGGAC ATTTGTATCT CATGTGGGAG
1701 CCTCAATGTC ACCCTGGAGC ACCCACTCTT CATTGGAGGC ATGTGCCAGA
1751 ACTGTAAGAA CTGCTTCTTG GAGTGTGCTT ACCAGTATGA CGACGATGGG
1801 TACCAGTCCT ATTGCACCAT CTGCTGTGGG GGGCGTGAAG TGCTCATGTG
1851 TGGGAACAAC AACTGCTGCA GGTGCTTTTG TGTCGAGTGT GTGGATCTCT
1901 TGGTGGGGCC AGGAGCTGCT CAGGCAGCCA TTAAGGAAGA CCCCTGGAAC
1951 TGCTACATGT GCGGGCATAA GGGCACCTAT GGGCTGCTGC GAAGACGGGA
2001 AGACTGGCCT TCTCGACTCC AGATGTTCTT TGCCAATAAC CATGACCAGG
2051 AATTTGACCC CCCAAAGGTT TACCCACCTG TGCCAGCTGA GAAGAGGAAG
2101 CCCATCCGCG TGCTGTCTCT CTTTGATGGG ATTGCTACAG GGCTCCTGGT
2151 GCTGAAGGAC CTGGGCATCC AAGTGGACCG CTACATTGCC TCCGAGGTGT
2201 GTGAGGACTC CATCAGGTG GGCATGGTGC GGCACCAGGG AAAGATCATG
2251 TACGTGGGGG ACGTCCGCAG CGTCACACAG AAGCATATCC AGGAGTGGGG
2301 CCCATTGCAC CTGGTGATTG GAGGCAGTCC CTGCAATGAC CTCTCCATTG

FIG. 1A-2

2351 TCAACCCTGC CCGCAAGGGA CTTTATGAGG GTACTGGCCG CCTCTTCTTT
2401 GAGTTCTACC GCCTCCTGCA TGATGCGCGG CCCAAGGAGG GAGATGATCG
2451 CCCCTTCTTC TGGCTCTTTG AGAATGTGGT GGCCATGGGC GTTAGTGACA
2501 AGAGGGACAT CTCGCGATTT CTTGAGTCTA ACCCCGTGAT GATTGACGCC
2551 AAAGAAGTGT CTGCTGCACA CAGGGCCCGT TACTTCTGGG GTAACCTTCC
2601 TGGCATGAAC AGGCCTTTGG CATCCACTGT GAATGATAAG CTGGAGCTGC
2651 AAGAGTGTCT GGAGCACGGC AGAATAGCCA AGTTCAGCAA AGTGAGGACC
2701 ATTACCACCA GGTCAAATC TATAAAGCAG GGCAAAGACC AGCATTTCCC
2751 CGTCTTCATG AACGAGAAGG AGGACATCCT GTGGTGCACT GAAATGAA
2801 GGGTGTTTGG CTTCCCGTC CACTACACAG ACGTCTCAA CATGAGCCGC
2851 TTGGCGAGGC AGAGACTGCT GGGCCGATCG TGGAGCGTGC CGGTCATCCG
2901 CCACCTCTTC GCTCCGCTGA AGGAATATTT TGCTTGTTG TAAGGGACAT
2951 GGGGGCAAAC TGAAGTAGTG ATGATAAAAA AGTTAAACAA ACAAACAAAC
3001 AAAAAACAAA ACAAACAAT AAAACACCAA GAACGAGAGG ACGGAGAAAA
3051 GTTCAGCACC CAGAAGAGAA AAAGGAATTT AAAGCAAACC ACAGAGGAGG
3101 AAAACGCCGG AGGGCTTGGC CTTGCAAAAG GGTGGACAT CATCTCCTGA
3151 GTTTTCAATG TTAACCTTCA GTCCTATCTA AAAAGCAAAA TAGGCCCCTC
3201 CCCTTCTTCC CCTCCGGTCC TAGGAGGCGA ACTTTTGTG TTCTACTCTT
3251 TTTCAGAGGG GTTTTCTGTT TGTTTGGGTT TTTGTTTCTT GCTGTGACTG
3301 AAACAAGAGA GTTATTGCAG CAAAATCAGT AACAACAAAA AGTAGAAATG
3351 CCTTGGAGAG GAAAGGGAGA GAGGGAAAAT TCTATAAAAA CTTAAAATAT
3401 TGGTTTTTTT TTTTTTCTT TTTCTATATA TCTCTTTGGT TGTCTCTAGC
3451 CTGATCAGAT AGGAGCACAA ACAGGAAGAG AATAGAGACC CTCGGAGGCA
3501 GAGTCTCCTC TCCCACCCCC CGAGCAGTCT CAACAGCACC ATTCCTGGTC

FIG. 1A-3

3551 ATGCAAAACA GAACCCAAC AGCAGCAGGG CGCTGAGAGA ACACCACACC
3601 AGACACTTTC TACAGTATTT CAGGTGCCTA CCACACAGGA AACCTGAAG
3651 AAAACCAGTT TCTAGAAGCC GCTGTACCT CTGTGTTACA GTTTATATAT
3701 ATATGATAGA TATGAGATAT ATATATATAA AAGGTACTGT TAACTACTGT
3751 ACATCCCGAC TTCATAATGG TGCTTTCAA ACAGCGAGAT GAGCAAAGAC
3801 ATCAGCTTCC GCCTGGCCCT CTGTGCAAAG GGTTCAGCC CAGGATGGGG
3851 AGAGGGGAGC AGCTGGAGGG GGTTTTAAAC AACTGAAGGA TGACCCATAT
3901 CACCCCCCAC CCCTGCCCCA TGCCTAGCTT CACCTGCCAA AAAGGGGCTC
3951 AGCTGAGGTG GTCGGACCCT GGGGAAGCTG AGTGTGGAAT TTATCCAGAC
4001 TCGCGTGCAA TAACCTTAGA ATATGAATCT AAAATGACTG CCTCAGAAAA
4051 ATGGCTTGAG AAAACATTGT CCCTGATTTT GAATTCGTCA GCCACGTTGA
4101 AGCCCCCTTG TGGATCAGA AATATTCCAG AGTGAGGGAA AGTGACCCGC
4151 CATTAAACCC NCCTGGAGCA AATAAAAAAA CATACAAAAT GT

FIG. 1A-4

Mouse Dnmt3b1 DNA Sequence

1 GAATTCCGGG CGCCGGGGTT AAGCGGCCCA AGTAAACGTA GCGCAGCGAT
51 CGGCGCCGGA GATTGCGGAA CCCGACACTC CGCGCCGCCC GCCGGCCAGG
101 ACCCGCGGCG CGATCGCGGC GCCGCGCTAC AGCCAGCCTC ACGACAGGCC
151 CGCTGAGGCT TGTGCCAGAC CTTGGAAACC TCAGGTATAT ACCTTTCCAG
201 ACGCGGGATC TCCCCTCCCC CATCCATAGT GCCTTGGGAC CAAATCCAGG
251 GCCTTCTTTC AGGAAACAAT GAAGGGAGAC AGCAGACATC TGAATGAAGA
301 AGAGGGTGCC AGCGGGTATG AGGAGTGCAT TATCGTTAAT GGGAACTTCA
351 GTGACCACTC CTCAGACACG AAGGATGCTC CCTCACCCCC AGTCTTGGAG
401 GCAATCTGCA CAGAGCCAGT CTGCACACCA GAGACCAGAG GCCGCAGGTC
451 AAGTCCCGG CTGTCTAAGA GGGAGGTCTC CAGCCTTCTG AATTACACGC
501 AGGACATGAC AGGAGATGGA GACAGAGATG ATGAAGTAGA TGATGGGAAT
551 GGCTCTGATA TTCTAATGCC AAAGCTCACC CGTGAGACCA AGGACACCAG
601 GACGCGCTCT GAAAGCCCGG CTGTCCGAAC CCGACATAGC AATGGGACCT
651 CCAGCTTGGA GAGGCAAAGA GCCTCCCCCA GAATCACCCG AGGTCCGCAG
701 GGCCGCCACC ATGTGCAGGA GTACCCTGTG GAGTTTCCGG CTACCAGGTC
751 TCGGAGACGT CGAGCATCGT CTTAGCAAG CAGCCATGG TCATCCCCTG
801 CCAGCGTCGA CTTATGGA GAAGTGACAC CTAAGAGCGT CAGTACCCCA
851 TCAGTTGACT TGAGCCAGGA TGGAGATCAG GAGGGTATGG ATACCACACA
901 GGTGGATGCA GAGAGCAGAG ATGGAGACAG CACAGAGTAT CAGGATGATA
951 AAGAGTTTGG AATAGGTGAC CTCGTGTGGG GAAAGATCAA GGGCTTCTCC
1001 TGGTGGCCTG CCATGGTGGT GTCCTGGAAA GCCACCTCCA AGCGACAGGC

FIG. 1B-1

1051 CATGCCCGGA ATGCCGCTGGG TACAGTGGTT TGGTGATGGC AAGTTTTCTG
1101 AGATCTCTGC TGACAAACTG GTGGCTCTGG GGCTGTTTCTG CCAGCACTTT
1151 AATCTGGCTA CCTTCAATAA GCTGGTTTCT TATAGGAAGG CCATGTACCA
1201 CACTCTGGAG AAAGCCAGGG TTCGAGCTGG CAAGACCTTC TCCAGCAGTC
1251 CTGGAGAGTC ACTGGAGGAC CAGCTGAAGC CCATGCTGGA GTGGGCCCAC
1301 GGTGGCTTCA AGCCTACTGG GATCGAGGGC CTCAAACCCA ACAAGAAGCA
1351 ACCAGTGGTT AATAAGTCGA AGGTGCGTCG TTCAGACAGT AGGAACTTAG
1401 AACCAGGAG ACGCGAGAAC AAAAGTCGAA GACGCACAAC CAATGACTCT
1451 GCTGCTTCTG AGTCCCCCCC ACCCAAGCGC CTCAAGACAA ATAGCTATGG
1501 CGGGAAGGAC CGAGGGGAGG ATGAGGAGAG CCGAGAACCG ATGGCTTCTG
1551 AAGTCACCAA CAACAAGGGC AATCTGGAAG ACCGCTGTTT GTCCTGTGGA
1601 AAGAAGAACC CTGTGTCCTT CCACCCCCTC TTTGAGGGTG GGCTCTGTCA
1651 GAGTTGCCGG GATCGCTTCC TAGAGCTCTT CTACATGTAT GATGAGGACG
1701 GCTATCAGTC CTAAGTCACC GTGTGCTGTG AGGGCCGTGA ACTGCTGCTG
1751 TGCAGTAACA CAAGCTGCTG CAGATGCTTC TGTGTGGAGT GTCTGGAGGT
1801 GCTGGTGGGC GCAGGCACAG CTGAGGATGC CAAGCTGCAG GAACCCTGGA
1851 GCTGCTATAT GTGCCTCCCT CAGCGCTGCC ATGGGGTCCT CCGACGCAGG
1901 AAAGATTGGA ACATGCGCCT GCAAGACTTC TTCCTACTG ATCCTGACCT
1951 GGAAGAATTT GAGCCACCCA AGTTGTACCC AGCAATTCCT GCAGCCAAAA
2001 GGAGGCCCAT TAGAGTCCTG TCTCTGTTTG ATGGAATTGC AACGGGGTAC
2051 TTGGTGCTCA AGGAGTTGGG TATTAAAGTG GAAAAGTACA TTGCCTCCGA
2101 AGTCTGTGCA GAGTCCATCG CTGTGGGAAC TGTTAAGCAT GAAGGCCAGA
2151 TCAAATATGT CAATGACGTC CGGAAAATCA CCAAGAAAAA TATTGAAGAG
2201 TGGGGCCCGT TCGACTTGGT GATTGGTGGA AGCCCATGCA ATGATCTCTC

FIG. 1B-2

2251 TAACGTCAAT CCTGCCCGCA AAGGTTTATA TGAGGGCACA GGAAGGCTCT
2301 TCTTCGAGTT TTACCACTTG CTGAATTATA CCCGCCCCAA GGAGGGCGAC
2351 AACCGTCCAT TCTTCTGGAT GTTCGAGAAT GTTGTGGCCA TGAAAGTGAA
2401 TGACAAGAAA GACATCTCAA GATTCCTGGC ATGTAACCCA GTGATGATCG
2451 ATGCCATCAA GGTGTCTGCT GCTCACAGGG CCCGGTACTT CTGGGGTAAC
2501 CTACCCGGAA TGAACAGGCC CGTGATGGCT TCAAAGAATG ATAAGCTCGA
2551 GCTGCAGGAC TGCCTGGAGT TCAGTAGGAC AGCAAAGTTA AAGAAAGTGC
2601 AGACAATAAC CACCAAGTCG AACTCCATCA GACAGGGCAA AAACCAGCTT
2651 TTCCCTGTAG TCATGAATGG CAAGGACGAC GTTTTGTGGT GCACTGAGCT
2701 CGAAAGGATC TTCGGCTTCC CTGCTACTA CACGGACGTG TCCAACATGG
2751 GCCGCGGGCC CCGTCAGAAG CTGCTGGGCA GGTCTGGAG TGTACCGGTC
2801 ATCAGACACC TGTTTGCCCC CTTGAAGGAC TACTTTGCCT GTGAATAGTT
2851 CTACCCAGGA CTGGGGAGCT CTCGGTCAGA GCCAGTGCCC AGAGTCACCC
2901 CTCCCTGAAG GCACCTCACC TGTCCCTTT TTAGCTCACC TGTGTGGGGC
2951 CTCACATCAC TGTACCTCAG CTTTCTCCTG CTCAGTGGGA GCAGAGCCTC
3001 CTGGCCCTTG CAGGGGAGCC CCGGTGCTCC CTCCGTGTGC ACAGCTCAGA
3051 CCTGGCTGCT TAGAGTAGCC CGGCATGGTG CTCATGTTCT CTTACCCTGA
3101 AACTTTAAAA CTTGAAGTAG GTAGTAAGAT GGCTTTCTTT TACCCTCCTG
3151 AGTTTATCAC TCAGAAGTGA TGGCTAAGAT ACCAAAAAAA CAAACAAAAA
3201 CAGAAACAAA AAACAAAAAA AAACCTCAAC AGCTCTCTTA GTACTCAGGT
3251 TCATGCTGCA AAATCACTTG AGATTTTGTT TTTAAGTAAC CCGTGCTCCA
3301 CATTTGCTGG AGGATGCTAT TGTGAATGTG GGCTCAGATG AGCAAGGTCA
3351 AGGGGCCAAA AAAAATTCCC CCTCTCCCC CAGGAGTATT TGAAGATGAT
3401 GTTTATGGTT TAAGTCTTCC TGGCACCTTC CCCTTGCTTT GGTACAAGGG

FIG. 1B-3

3451 CTGAAGTCCT GTTGGTCTTG TAGCATTTC CAGGATGATG ATGTCAGCAG
3501 GGATGACATC ACCACCTTTA GGGCTTTTCC CTGGCAGGGG CCCATGTGGC
3551 TAGTCCTCAC GAAGACTGGA GTAGAATGTT TGGAGCTCAG GAAGGGTGGG
3601 TGGAGTGGCC CTCTTCCAGG TGTGAGGGAT ACGAAGGAGG AAGCTTAGGG
3651 AAATCCATTC CCCACTCCCT CTTGCCAAAT GAGGGGCCCA GTCCCCAACA
3701 GCTCAGGTCC CCAGAACCCC CTAGTTCTC ATGAGAAGCT AGGACCAGAA
3751 GCACATCGTT CCCCTTATCT GAGCAGTGTT TGGGGAAC TA CAGTGA AAAC
3801 CTTCTGGAGA TGTTAAAAGC TTTTACCCC ACGATAGATT GTGTTTTTAA
3851 GGGGTGCTTT TTTTAGGGGC ATCACTGGAG ATAAGAAAGC TGCATTTTCA
3901 AAATGCCATC GTAATGGTTT TTAACACCT TTTACCTAAT TACAGGTGCT
3951 ATTTTATAGA AGCAGACAAC ACTTCTTTTT ATGACTCTCA GACTTCTATT
4001 TTCATGTTAC CATTTTTTTT GTAAC TCGCA AGGTGTGGGC TTTTGTAAC T
4051 TCACAGGTGT GGGGAGAGAC TGCCTTGTTT CAACAGTTTG TCTCCACTGG
4101 TTTCTAATTT TTAGGTGCAA AGATGACAGA TGCC CAGAGT TTACCTTTCT
4151 GGTGATTAA AGTTGTATTT CTCTAAAAA AAAAAAAAAA AAAAA

FIG. 1B-4

Human DNMT3A DNA Sequence

1 GCCGCGG CACCAGGGCG CGCAGCCGGG
28 CCGGCCCCGAC CCCACCGGCC ATACGGTGA GCCATCGAAG CCCCCACCCA
78 CAGGCTGACA GAGGCACCGT TCACCAGAGG GCTCAACACC GGGATCTATG
128 TTTAAGTTTT AACTCTCGCC TCCAAAGACC ACGATAATTC CTTCCCCAAA
178 GCCCAGCAGC CCCCCAGCCC CGCGCAGCCC CAGCCTGCCT CCCGGCGCCC
228 AGATGCCCCG CATGCCCTCC AGCGGCCCCG GGGACACCAG CAGCTCTGCT
278 GCGGAGCGGG AGGAGGACCG AAAGGACGGA GAGGAGCAGG AGGAGCCGCG
328 TGGCAAGGAG GAGCGCCAAG AGCCCAGCAC CACGGCACGG AAGGTGGGGC
378 GGCCTGGGAG GAAGCGCAAG CACCCCCCGG TGGAAAGCGG TGACAGGCCA
428 AAGGACCCTG CGGTGATCTC CAAGTCCCCA TCCATGGCCC AGGACTCAGG
478 CGCCTCAGAG CTATTACCCA ATGGGGACTT GGAGAAGCGG AGTGAGCCCC
528 AGCCAGAGGA GGGGAGCCCT GCTGGGGGGC AGAAGGGCGG GGGCCCAGCA
578 GAGGGAGAGG GTGCAGCTGA GACCCTGCCT GAAGCCTCAA GAGCAGTGA
628 AAATGGCTGC TGCACCCCCA AGGAGGGCCG AGGAGCCCCT GCAGAAGCGG
678 GCAAAGAACA GAAGGAGACC AACATCGAAT CCATGAAAAT GGAGGGCTCC
728 CGGGGCCCGC TCGGGGTGG CTTGGGCTGG GAGTCCAGCC TCCGTCAGCG
778 GCCCATGCCG AGGCTCACCT TCCAGGCGGG GGACCCCTAC TACATCAGCA
828 AGCGCAAGCG GGACGAGTGG CTGGCAGCT GGAAGGGA GGCTGAGAAG
878 AAAGCCAAGG TCAGTGCAGG AATGAATGCT GTGAAGAAA ACCAGGGGCC
928 CGGGGAGTCT CAGAAGGTGG AGGAGGCCAG CCCTCCTGCT GTGCAGCAGC
978 CCACTGACCC CGCATCCCC ACTGTGGCTA CCACGCCTGA GCCCGTGGGG
1028 TCCGATGCTG GGGACAAGAA TGCCACCAAA GCAGGCGATG ACGAGCCAGA

FIG. 1C-1

1078 GTACGAGGAC GGCCGGGGCT TTGGCATTGG GGAGCTGGTG TGGGGGAAAC
1128 TGCCGGGGCTT CTCCTGGTGG CCAGGCCGCA TTGTGTCTTG GTGGATGACC
1178 GGCCGGAGCC GAGCAGCTGA AGGCACCCGC TGGGTCATGT GGTTCGGAGA
1228 CGGCAAATTC TCAGTGGTGT GTGTTGAGAA GCTGATGCCG CTGAGCTCGT
1278 TTTGCAGTGC GTTCCACCAG GCCACGTACA ACAAGCAGCC CATGTACCGC
1328 AAAGCCATCT ACGAGGTCCT GCAGGTGGCC AGCAGCCGCG CGGGGAAGCT
1378 GTTCCCGGTG TGCCACGACA GCGATGAGAG TGACACTGCC AAGGCCGTGG
1428 AGGTGCAGAA CAAGCCCATG ATTGAATGGG CCCTGGGGGG CTTCCAGCCT
1478 TCTGGCCCTA AGGGCCTGGA GCCACCAGAA GAAGAGAAGA ATCCCTACAA
1528 AGAAGTGAC ACGGACATGT GGGTGAACC TGAGGCAGCT GCCTACGCAC
1578 CACCTCCACC AGCCAAAAAG CCCCAGGAAGA GCACAGCGGA GAAGCCCAAG
1628 GTCAAGGAGA TTATTGATGA GCGCACAAGA GAGCGGCTGG TGTACGAGGT
1678 GCGGCAGAAG TGCCGGAACA TTGAGGACAT CTGCATCTCC TGTGGGAGCC
1728 TCAATGTTAC CCTGGAACAC CCCCTCTTCG TTGGAGGAAT GTGCCAAAAC
1778 TGCAAGAACT GCTTTCTGGA GTGTGCGTAC CAGTACGACG ACGACGGCTA
1828 CCAGTCCTAC TGCACCATCT GCTGTGGGGG CCGTGAGGTG CTCATGTGCG
1878 GAAACAACAA CTGCTGCAGG TGCTTTTGGC TGGAGTGTGT GGACCTCTTG
1928 GTGGGGCCCG GGGCTGCCCC GGCAGCCATT AAGGAAGACC CCTGGAAGTG
1978 CTACATGTGC GGGCACAAGG GTACCTACGG GCTGCTGCGG CGGCGAGAGG
2028 ACTGGCCCTC CCGGCTCCAG ATGTTCTTCG CTAATAACCA CGACCAGGAA
2078 TTTGACCCTC CAAAGGTTTA CCCACCTGTC CCAGCTGAGA AGAGGAAGCC
2128 CATCCGGGTG CTGTCTCTCT TTGATGGAAT CGCTACAGGG CTCCTGGTGC
2178 TGAAGGACTT GGGCATTGAG GTGGACCGCT ACATTGCCTC GGAGGTGTGT

FIG. 1C-2

2228 GAGGACTCCA TCACGGTGGG CATGGTCCGG CACCAGGGGA AGATCATGTA
2278 CGTCGGGGAC GTCCGCAGCG TCACACAGAA GCATATCCAG GAGTGGGGCC
2328 CATTGATCT GGTGATTGGG GGCAGTCCCT GCAATGACCT CTCCATCGTC
2378 AACCTGCTC GCAAGGGCCT CTACGAGGGC ACTGGCCGGC TCTTCTTTGA
2428 GTTCTACCGC CTCCTGCATG ATGCCGGGCC CAAGGAGGGA GATGATCGCC
2478 CCTTCTTCTG GCTCTTTGAG AATGTGGTGG CCATGGGCGT TAGTGACAAG
2528 AGGGACATCT CGCGATTCT CGAGTCCAAC CCTGTGATGA TTGATGCCAA
2578 AGAAGTGTC GCTGCACACA GGGCCCGCTA CTTCTGGGGT AACCTTCCCG
2628 GTATGAACAG GCCGTTGGCA TCCACTGTGA ATGATAAGCT GGAGCTGCAG
2678 GAGTGTCTGG AGCATGGCAG GATAGCCAAG TTCAGCAAAG TGAGGACCAT
2728 TACTACGAGG TCAAACCTCA TAAAGCAGGG CAAAGACCAG CATTTTCTG
2778 TCTTCATGAA TGAGAAAGAG GACATCTTAT GGTGCACTGA AATGGAAAGG
2828 GTATTGGTT TCCAGTCCA CTATACTGAC GTCTCCAACA TGAGCCGCTT
2878 GCGGAGGCAG AGACTGCTGG GCCGGTCATG GAGCGTGCCA GTCATCCGCC
2928 ACCTCTTCGC TCCGCTGAAG GAGTATTTTG CGTGTGTGTA AGGGACATGG
2978 GGGCAAACCTG AGGTAGCGAC ACAAAGTTAA ACAAACAAAC AAAAAACACA
3028 AAACATAATA AAACACCAAG AACATGAGGA TGGAGAGAAG TATCAGCACC
3078 CAGAAGAGAA AAAGGAATTT AAAACAAAAA CCACAGAGGC GGAAATACCG
3128 GAGGGCTTTG CCTTGGGAAA AGGGTTGGAC ATCATCTCCT GATTTTTCAA
3178 TGTTATTCTT CAGTCCTATT TAAAAACAAA ACCAAGCTCC CTTCCCTTCC
3228 TCCCCCTTCC CTTTTTTTTT GGTGAGACCT TTTATTTTCT ACTCTTTTCA
3278 GAGGGGTTTT CTGTTTGTTT GGGTTTGTG TCTTGCTGTG ACTGAAACAA
3328 GAAGGTATT GCAGCAAAAA TCAGTAACAA AAAATAGTAA CAATACCTTG
3378 CAGAGGAAAG GTGGGAGGAG AGGAAAAAAG GGAAATTTTT AAAGAAATCT

FIG. 1C-3

3428 ATATATTGGG TTGTTTTTTT TTTTGTITTT TGTTTTTTTT TTTTGGGTTT
3478 TTTTTTTTTA CTATATATCT TTTTTTTGTT GTCTCTAGCC TGATCAGATA
3528 GGAGCACAAG CAGGGGACGG AAAGAGAGAG ACACTCAGGC GGCAGCATTC
3578 CCTCCCAGCC ACTGAGCTGT CGTGCCAGCA CCATTCCTGG TCACGCAAAA
3628 CAGAACCCAG TTAGCAGCAG GGAGACGAGA ACACCACACA AGACATTTTT
3678 CTACAGTATT TCAGGTGCCT ACCACACAGG AAACCTTGAA GAAAATCAGT
3728 TTCTAGAAGC CGCTGTTACC TCTTGTTTAC AGTTTATATA TATATGATAG
3778 ATATGAGATA TATATATAAA AGGTACTGTT AACTACTGTA CAACCCGACT
3828 TCATAATGGT GCTTTCAAAC AGCGAGATGA GTAAAAACAT CAGCTTCCAC
3878 GTTGCCTTCT GCGCAAAGGG TTTCACCAAG GATGGAGAAA GGGAGACAGC
3928 TTGCAGATGG CGCGTTCTCA CGGTGGGCTC TTCCCCTTGG TTTGTAACGA
3978 AGTGAAGGAG GAGAACTTGG GAGCCAGGTT CTCCTGCCA AAAAGGGGGC
4028 TAGATGAGGT GGTGGGGCCC GTGGACAGCT GAGAGTGGGA TTCATCCAGA
4078 CTCATGCAAT AACCCTTTGA TTGTTTTCTA AAAGGAGACT CCCTCGGCAA
4128 GATGGCAGAG GGTACGGAGT CTTAGGCCCC AGTTTCTCAC TTTAGCCAAT
4178 TCGAGGGCTC CTTGTGGTGG GATCAGAACT AATCCAGAGT GTGGGAAAGT
4228 GACAGTCAAA ACCCCACCTG GAGCAAATAA AAAACATAC AAAACGTAA
4278 AAAAAAAAAA AAAAAA

FIG. 1C-4

Human DNMT3B1 DNA Sequence:

1 GGCOCGAAT TCGGCACGAG CCCTGCACGG CCGCCAGCCG GCCTCCCGCC
51 AGCCAGCCCC GACCCGCGGC TCCGCCGCCC AGCCGCGCCC CAGCCAGCCC
101 TCGGCAGGA AAGCATGAAG GGAGACACCA GGCATCTCAA TGGAGAGGAG
151 GACGCCGGCG GGAGGGAAGA CTCGATCCTC GTCAACGGGG CCTGCAGCGA
201 CCAGTCCTCC GACTCGCCCC CAATCCTGGA GGCTATCCGC ACCCCGGAGA
251 TCAGAGGCCG AAGATCAAGC TCGCGACTCT CCAAGAGGGA GGTGTCCAGT
301 CTGCTAAGCT ACACACAGGA CTTGACAGGC GATGGCGACG GGAAGATGG
351 GGATGGCTCT GACACCCAG TCATGCCAAA GCTCTTCCG GAAACCAGGA
401 CTCGTTGAGA AAGCCCAGCT GTCCGAAGTC GAAATAACAA CAGTGTCTCC
451 AGCCGGGAGA GGCACAGGCC TTCCCCAGT TCCACCCGAG GCCGGCAGGG
501 CCGCAACCAT GTGGACGAGT CCCCCGTGGA GTTCCCGCT ACCAGGTCCC
551 TGAGACGGCG GGCAACAGCA TCGGCAGGAA CGCCATGGCC GTCCCTCCC
601 AGCTCTTACC TTACCATCGA CCTCACAGAC GACACAGAGG ACACACATGG
651 GACGCCCCAG AGCAGCAGTA CCCCCTACGC CCGCCTAGCC CAGGACAGCC
701 AGCAGGGGGG CATGGAGTCC CCGCAGGTGG AGGCAGACAG TGGAGATGGA
751 GACAGTTCAG AGTATCAGGA TGGGAAGGAG TTTGGAATAG GGGACCTCGT
801 GTGGGAAAG ATCAAGGGCT TCTCCTGGTG GCCCGCATG GTGGTGTCTT
851 GGAAGGCCAC CTCCAAGCGA CAGGCTATGT CTGGCATGCG GTGGGTCCAG
901 TGGTTTGGCG ATGGCAAGTT CTCCGAGGTC TCTGCAGACA AACTGGTGGC
951 ACTGGGGCTG TTCAGCCAGC ACTTTAATTT GGCCACCTTC AATAAGCTCG
1001 TCTCCTATCG AAAAGCCATG TACCATGCTC TGGAGAAAGC TAGGGTGCGA
1051 GCTGGCAAGA CCTTCCCCAG CAGCCCTGGA GACTCATTGG AGGACCAGCT
1101 GAAGCCCATG TTGGAGTGGG CCCACGGGGG CTTCAAGCCC ACTGGGATCG
1151 AGGGCCTCAA ACCCAACAAC ACGCAACCAG TGGTAATAA GTCGAAGGTG

FIG. 1D-1

1201 CGTCGTGCAG GCAGTAGGAA ATTAGAATCA AGGAAATACG AGAACAAGAC
1251 TCGAAGACGC ACAGCTGACC ACTCAGCCAC CTCTGACTAC TGCCCCGCAC
1301 CCAAGCCCTT CAAGACAAAT TGCTATAACA ACGGCAAAGA CCGAGGGGAT
1351 GAAGATCAGA GCCGAGAACA AATGGCTTCA GATGTTGCCA ACAACAAGAG
1401 CAGCCTGGAA GATGGCTGTT TGTCTTGTGG CAGGAAAAAC CCCGTGTCCT
1451 TCCACCCTCT CTTTGAGGGG GGGCTCTGTC AGACATGCCG GGATCGCTTC
1501 CTTGAGCTGT TTTACATGTA TGATGACGAT GGCTATCAGT CTTACTGCAC
1551 TGTGTGCTGC GAGGGCCGAG AGCTGCTGCT TTGCAGCAAC ACGAGCTGCT
1601 GCCGGTGTTT CTGTGTGGAG TGCCTGGAGG TGCTGGTGGG CACAGGCACA
1651 GCGGCCGAGG CCAAGCTTCA GGAGCCCTGG AGCTGCTACA TGTGTCTCCC
1701 GCAGCGCTGT CATGGCGTCC TGCGGCGCCG GAAGGACTGG AACGTGCGCC
1751 TGCAGGCCCTT CTTACCAGT GACACGGGGC TTGAATACGA AGCCCCCAAG
1801 CTGTACCCTG CCATTCCCGC AGCCCGAAGG CGGCCCATTC GAGTCCTGTC
1851 ATTGTTTGAT GGCATCGCGA CAGGCTACCT AGTCCTCAAA GAGTTGGGCA
1901 TAAAGGTAGG AAAGTACGTC GCTTCTGAAG TGTGTGAGGA GTCCATTGCT
1951 GTTGAACCG TGAAGCACGA GGGGAATATC AAATACGTGA ACGACGTGAG
2001 GAACATCACA AAGAAAAATA TTGAAGAATG GGGCCCATTT GACTTGGTGA
2051 TTGGCGGAAG CCCATGCAAC GATCTCTCAA ATGTGAATCC AGCCAGGAAA
2101 GGCCTGTATG AGGGTACAGG CCGGCTCTTC TTCGAATTTT ACCACCTGCT
2151 GAATTACTCA CGCCCCAAGG AGGGTGATGA CCGGCCGTTT TTCTGGATGT
2201 TTGAGAATGT TGTAGCCATG AAGGTTGGCG ACAAGAGGGA CATCTACCGG
2251 TTCCTGGAGT GTAATCCAGT GATGATTGAT GCCATCAAAG TTTCTGCTGC
2301 TCACAGGGCC CGATACTTCT GGGGCAACCT ACCCGGGATG AACAGGCCCC
2351 TGATAGCATC AAAGAATGAT AAACGAGC TGCAGGACTG CTTGGAATAC
2401 AATAGGATAG CCAAGTTAAA GAAAGTACAG ACAATAACCA CCAAGTCGAA

FIG. 1D-2

2451 CTCGATCAAA CAGGGGAAAA ACCAACTTTT CCCTGTTGTC ATGAATGGCA
 2501 AAGAAGATGT TTTGTGGTGC ACTGAGCTCG AAAGGATCTT TGGCTTTCCT
 2551 GTGCACTACA CAGACGTGTC CAACATGGGC CGTGGTGCCC GCCAGAAGCT
 2601 GCTGGGAAGG TCCTGGAGCG TGCCTGTCAT CCGACACCTC TTCGCCCCCTC
 2651 TGAAGGACTA CTTTGCATGT GAATAGTTCC AGCCAGGCCC CAAGCCCCACT
 2701 GGGGTGTGTG GCAGAGCCAG GACCCAGGAG GTGTGATTCC TGAAGGCATC
 2751 CCCAGGCCCT GCTCTTCCTC AGCTGTGTGG GTCATACCGT GTACCTCAGT
 2801 TCCCTCTTGC TCAGTGGGGG CAGAGCCACC TGACTCTTGC AGGGGTAGCC
 2851 TGAGGTGCCG CCTCCTTGTG CACAAATCAG ACCTGGCTGC TTGGAGCAGC
 2901 CTAACACGGT GCTCATTTTT TCTTCTCCTA AAACTTTAAA ACTTGAAGTA
 2951 GGTAGCAACG TGGCTTTTTT TTTTCCCTT CCTGGGTCTA CCACTCAGAG
 3001 AAACAATGGC TAAGATACCA AAACCACAGT GCCGACAGCT CTCCAATACT
 3051 CAGGTTAATG CTGAAAAATC ATCCAAGACA GTTATTGCAA GAGTTTAATT
 3101 TTTGAAAAC TGGTACTGCT ATGTGTTTAC AGACGTGTGC AGTTGTAGGC
 3151 ATGTAGCTAC AGGACATTTT TAAGGGCCCA GGATCGTTTT TTCCAGGGC
 3201 AAGCAGAAGA GAAAATGTTG TATATGTCTT TTACCCGGCA CATTCCCCTT
 3251 GCCTAAATAC AAGGGCTGGA GTCTGCACGG GACCTATTAG AGTATTTTCC
 3301 ACAATGATGA TGATTTACGC AGGGATGACG TCATCATCAC ATTCAGGGCT
 3351 ATTTTTTCCC CCACAAACCC AAGGGCAGGG GCCACTCTTA GCTAAATCCC
 3401 TCCCCGTGAC TGCAATAGAA CCCTCTGGGG AGCTCAGGAA GGGGTGTGCT
 3451 GAGTTCTATA ATATAAGCTG CCATATATTT TGTAGACAAG TATGGCTCCT
 3501 CCATATCTCC CTCTTCCCTA GGAGAGGAGT GTGAAGCAAG GAGCTTAGAT
 3551 AAGACACCCC CTCAAACCCA TTCCCTCTCC AGGAGACCTA CCCTCCACAG
 3601 GCACAGGTCC CCAGATGAGA AGTCTGCTAC CCTCATTCTT CATCTTTTAA
 3651 CTAAACTCAG AGGCAGTGAC AGCAGTCAGG GACAGACATA CATTCTCAT

FIG. 1D-3

3701 ACCTTCCCCA CATCTGAGAG ATGACAGGGA AACTGCAAA GCTCGGTGCT
3751 CCCTTTGGAG ATTTTTTAAT CCTTTTTTAT TCCATAAGAA GTCGTTTTTA
3801 GGGAGAACGG GAATTCAGAC AAGCTGCATT TCAGAAATGC TGTCAATG
3851 GTTTTAAACA CCTTTTACTC TTCTTACTGG TGCTATTTTG TAGAATAAGG
3901 AACACGTTG ACAAGTTTTG TGGGGCTTTT TATACACTTT TTAAAATCTC
3951 AACTTCTAT TTTTATGTTT AACGTTTCA TTAAAATTTT TTTGTAAGTG
4001 GAGCCACGAC GTAACAAATA TCGGGAAAAA ACTGTGCCTT GTTTCACAG
4051 TTTTGTCTAA TTTTAGGCT GAAAGATGAC GGATGCCTAG AGTTTACCTT
4101 ATGTTTAATT AAAATCAGTA TTTGTCTAAA AAAAAAAAAA AAAAA

FIG. 1D-4

Mouse Dnmt3a Protein

1 MPSSGPGDTS SSSLEREDDR KEGEEQEENR GKEERQEPSA TARKVGRPGR
51 KRKHPPVESS DTPKDPVTT KSQPMQDSC PSDLLPNGDL EKRSEPPQEE
101 GSPAAGQKGG APAEGEGTET PPEASRAVEN GCCVTKEGRG ASAGEGKEQK
151 QTNIESMKME GSRGRLRGGL GWESSLRQRP MPRLTFQAGD PYYISKRRKD
201 EWLARWKREA EKKAKVIAVM NAVEENQASC ESQKVEEASP PAVQQPTDPA
251 SPTVATTPEP VGGDAGDKNA TKAADDEPEY EDGRGFGIGE LVWGKLRGFS
301 WWPGRIVSWW MTGRSRAAEG TRWVMFQDG KFSVVCVEKL MPLSSFCSAF
351 HQATYNQPM YRKAIYEVLO VASSRAGKLF PACHDSDESD SGKAVEVQNK
401 QMIEWALGGF QPSGPKGLEP PEEKNPYKE VYTDMMVEPE AAAYAPPPPA
451 KKPRKSTTEK PKVKEIIDER TRERLVYEVN QKCRNIEDIC ISCGSLNVTL
501 EHPLFIGGMC QNCKNCFLEC AYQYDDDGYY SYCTICCGGR EVLMCGNNNC
551 CRCFCVECVD LLVGPGAAQA AIKEDPWNCY MCGHKGTYGL LRRREDWPSR
601 LQMFFANNHD QEFDPPKVYP PVPAEKRKPI RVLSLFDGIA TGLLVKDLG
651 IQVDRIASE VCEDSITVGM VRHQGKIMYV GDVRSVTQKH IQEWGPFDLV
701 IGGSPCNDLS IVNPARKGLY EGTGRLFFEF YRLLHDARPK EGDDRPFFWL
751 FENVVAMGVS DKRDISRFLE SNPVMIDAKE VSAAHRARYF WGNLPGMNRP
801 LASTVNDKLE LQECLEHGRI AKFSKVRTIT TRSNSIKQKQ DQHFPVFMNE
851 KEDILWCTEM ERVFGFPVHY TDVSNMSRLA RQRLLGRSWS VPVIRHLFAP
901 LKEYFACV*

FIG. 2A

Mouse Dnmt3b1 Protein

1 MKGDSRHLNE EEGASGYEEC IIVNGNFSQD SSDTKDAPSP PVLEAICTEP
51 VCTPETRGRR SSSRLSKREV SSLLNYTQDM TGDGDRDDEV DDGNGSDILM
101 PKLTRETKDT RTRSESPAVR TRHSNGTSSL ERQRASPRIT RGRQGRHHVQ
151 EYPVEFPATR SRRRRASSSA STPWSSPASV DFMEEVTPKS VSTPSVDLSQ
201 DGDQEGMDTT QVDAESRDGD STEYQDDKEF GIGDLVWCKI KGFSWWPAMV
251 VSWKATSKRQ AMPGMRWWQW FGDGKFSEIS ADKLVALGLF SQHFNLATFN
301 KLVSYRKAMY HTLEKARVRA GKTFS SSPGE SLEDQLK PML EWAHGGFKPT
351 GIEGLKPNKK QPVVNKSKVR RSDSRNLEPR RRENKSRRRT TNSAASESP
401 PPKRLKTNSY GKGDRGEDEE SRERMASEVT NNKGNLEDRC LSCGKKNPVS
451 FHPLFEGGLC QSCRDRFLEL FYMYDEG YQ SYCTVCCEGR ELLLCNTSC
501 CRCFCVECLE VLVGAGTAED AKLQEPWSCY MCLPQRCHGV LRRRKDWNMR
551 LQDFFTTDPD LEEFEPKLY PAIPAAKRRP IRVLSLFDGI ATGYLVKEL
601 GIKVEKYIAS EVCAESIAVG TVKHEGQIKY VNDVRKITKK NIEEWGPFDL
651 VIGGSPCNDL SNVNPARKGL YEGTGR LFFE FYHLLNYTRP KEGDNRPF FW
701 MFENVVAMKV NDKKDISRFL ACNPVMIDAI KVSAAHRARY FWGNLPGMNR
751 PVMASKNDKL ELQDCLEFSR TAKLKKVQTI TTKSNSIRQG KNQLFPVVMN
801 GKDDVLWCTE LERIFGFPAH YTDVSNMGRG ARQKLLGRSW SVPVIRHLFA
851 PLKDYFACE*

FIG. 2B

Human DNMT3A Protein

1 MPAMPSSGPG DTSSSAAERE EDRKDGESEQE EPRGKEERQE PSTTARKVGR
51 PGRKRKHPPV ESGDTPKDPA VISKSPSMAQ DSGASELLPN GDLEKRSEPO
101 PEEGSPAGGQ KGGAPAELEG AAETLPEASR AVENGCCTPK EGRGAPAEAG
151 KEQKETNIES MKMEGSRGRL RGGLGWESSL RQRPMPLRTF QAGDPYYISK
201 RKRDEWLARW KREAEEKKAV IAGMNAVEEN QGPGESQKVE EASPPAVQQP
251 TDPASPTVAT TPEPVGSDAG DKNATKAGDD EPEYEDGRGF GIGELVWGKL
301 RGFSWWPGRI VSWMMTGRSR AAEGTRWVMW FGDGKFSVVC VEKLMPLSSF
351 CSAFHQATYN KQPMYRKAIY EVLQVASSRA GKLFPVCHDS DESDTAKAVE
401 VQNKPMIEWA LGGFQPSGPK GLEPPEEEKN PYKEVYTDMM VEPEAAAYAP
451 PPPAKKPRKS TAEKPKVKEI IDERTRERLV YEVRQKCRNI EDICISCGSL
501 NVTLEHPLFV GGMCQNCKNC FLECAQYDD DGYQSYCTIC CGGREVLMOG
551 NNNCCRCFCV ECVDLLVGPG AAQAAIKEDP WNCYMOGHKG TYGLLRRED
601 WPSRLQMFFA NNHDQEFDPP KVYPPVPAEK RKPVRVLSLF DGIATGLLV
651 KDLGIQVDRY IASEVCEDSI TVGMVRHQGK IMYVGDVRSV TQKHIEWGP
701 FDLVIGGSPC NDLSIVNPAR KGLYEGTGRL FFEFYRLLHD ARPKEGDDRP
751 FFWLFENVVA MGVS DKRDIS RFLESNPVMI DAKEVSAHR ARYFWGNLPG
801 MNRPLASTVN DKLELQECLE HGRIAKFSKV RTITTRSNSI KQKQDQHPV
851 FMNEKEDILW CTEMERVFGF PVHYTDVSNM SRLARQRLG RSWSPVIRH
901 LFAPLKEYFA CV*

FIG. 2C

Human DNMT3B1 Protein

1 MKGDTRHLNG EEDAGGREDS ILVNGACSDQ SSDSPPILEA IRTPEIRGRR
51 SSSRLSKREV SSLLSYTQDL TGDGDGEDGD GSOTPVMPKL FRETRTRSES
101 PAVRTRNNNS VSSRERHRPS PRSTRGRQGR NHVDESPVEF PATRSLRRRA
151 TASAGTPWPS PPSSYLTI DL TDDTEDTHGT PQSSSTPYAR LAQDSQQGGM
201 ESPQVEADSG DGDSEYQDG KEFGIGDLVW GKIKGFSWMP AMVVSWKATS
251 KRQAMSGMRW VQWFGDGKFS EVSADKLVAL GLFSQHFNLA TFNKLVSYRK
301 AMYHALEKAR VRAGKTFPSS PGDSLEDQLK PMLEWAHGGF KPTGIEGLKP
351 NNTQPVVNKS KVRRA GSRKL ESRKYENKTR RRTADDSATS DYCPAPKRLK
401 TNCYNGKDR GDEDQSREQM ASDVANNKSS LEDGCLSCGR KNPVSFHPLF
451 EGGLCQTCRD RFLELFYMYD DDGYQSYCTV CCEGRELLLC SNTSCCRFC
501 VECLEVLVGT GTAAEAKLQE PWSCYMCLPQ RCHGVLRRRK DWNVRLQAFF
551 TSDTGLEYEA PKLYPAIPAA RRRPIRVLSL FDGIATGYLV LKELGIKVGK
601 YVASEVCEES IAVGTVKHEG NIKYVNDVRN ITKKNIEEWG PFDLVIGGSP
651 CNDLSNVNPA RKGLYEGTGR LFFEFYHLLN YSRPKEGDDR PFFWMFENVV
701 AMKVGDKRDI SRFLECNPVM IDAIKVSAAH RARYFWGNLP GMNRPVIASK
751 NDKLELQDCL EYNRIAKLKK VQTITTKSNS IKQGNQLFP VMNGKEDVL
801 WCTELERIFG FPVHYTDVSN MGRGARQKLL GRSWSVPVIR HLFAPLKDYF
851 ACE*

FIG. 2D

Dnmt3a 1 MPSSGPGDTSSSSLEREDDRKEGEEQEENRGKEERQEPSATARKVGRPGR 50
 Dnmt3a 51 KRKHPPVESSDTPKDPVTTKSQPMQDSGPSD....LLPNGDLEKRSEP 96
 Dnmt3b 1MKGDSRHLNEEGASGYEECIIVNGNFSQSSD 33
 Dnmt3a 97 QPEEGSP....AAGQKGGAPAEGETTETPEAS.RAVENGCCVTKE..GR 139
 Dnmt3b 34 TKDAPSPPVLEAICTEPVCTPETRGRSSSRLSKREVSSLLNYTQDMTGD 83
 Dnmt3a 140 G.....ASAGEG.....KEQKQTNIESMKMEGSRGRLRGGLGWESSLRQ 178
 Dnmt3b 84 GDRDDEVDDGNGSDILMPKLTRETKDTRTRSESPAVRTRHSNGTSSLERQ 133
 Dnmt3a 179 RMPRLTFQAGDPYYISKRRDEWLARWKREAEEKAKVIAVMNAVEENQA 228
 Dnmt3b 134 RASPRITRGRQGRHHV.....QEYPVEFPATRSRRRRASSSASTPWSSPA 178
 Dnmt3a 229 SGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP 278
 Dnmt3b 179 SVDF..MEEVTPKSVSTP....SVDLSQDGDQEGMDTTQVDAESRDGDST 222
 Dnmt3a 279 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFG 328
 Dnmt3b 223 EYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMPGRWVQWFG 272
 Dnmt3a 329 DGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEV LQVASSRAGK 378
 Dnmt3b 273 DGKFSEISADKLVALGLFSQHFNLATFNKLVSRYKAMYHTLEKARVRAGK 322
 Dnmt3a 379 LFPACHDSDES DSGKAVEVQNQMIEWALGGFQPSGPKGLEPPEEEK..N 426
 Dnmt3b 323 TF.....SSSPGESLEDQLKPMLEWAHGGFKPTGIEGLKPNKKQPVVN 365
 Dnmt3a 427 PYKEVYTDMW.VEP.....EAAAYAPPPAKKPRKSTTEKPK 462
 Dnmt3b 366 KSKVRRSDSRNLEPRRRENKSRRTTND SAASESPPPKRLKTN SYGGKDR 415

FIG.3A-1

Dnmt3a 463 VKEIIDERTRERLVYEVQRKCRNIEDICISCGSLNVTLEHPFFIGGMCQN 512
 . || .|||: || |::| |::| | || |::|
 Dnmt3b 416 GE...DEESRERMASEVTNNKGNLEDRLSCGKKNPVSFHLPLFEGGLCQS 462
 Dnmt3a 513 CKNCFLECAQQYDDDDGYQSYCTICCGGREVLNCGNNCCRCFCVECVDLL 562
 |:: ||| | ||:|||||||::| |||.::| | .|||||||:::|
 Dnmt3b 463 CRDRFLELFYMYDEDGYQSYCTVCCEGRELLCSNTSCRCFCVECLEVL 512
 Dnmt3a 563 VGPGAAQAAIKEDPWNCYMGHKGTYGLLRREDWPSRLQMFFANNHD.Q 611
 || | |:: | ::||.|||| . :|.||||.|| ||| || . |:
 Dnmt3b 513 VGAGTAEDAKLQEPWSCYMCLPQRCHGVLRRRKDWNMRLQDFFTTDPDLE 562
 Dnmt3a 612 EFDPPKVYPPVPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRIYASEV 661
 ||:||||.|| :|| ||:||||||| ||||:||||.||:||||||
 Dnmt3b 563 EFEPKLYPAIPAAKRRPIRVLSLFDGIATGYLVKELGIKVEKYIASEV 612
 Dnmt3a 662 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSI 711
 | :|| || |::|. | || ||| :|.|. |:||||||| |||||
 Dnmt3b 613 CAESIAVGTVKHEGQIKYVNDVRKITKKNIEEWGPFDLVIGGSPCNDLSN 662
 Dnmt3a 712 VNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFWLFENVVAMGVSD 761
 ||||| ||||| ||||| ||. |||||.||||:||||| |. |
 Dnmt3b 663 VNPARKGLYEGTGRLFFEFYHLLNYTRPKEGDNRPFWMFENVVAMKVND 712
 Dnmt3a 762 KRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLEL 811
 |:||||| ||||| .||||||| ||||| . . |||||
 Dnmt3b 713 KKDISRFLACNPVMIDAIVSAHRARYFWGNLPGMNRPMASKNDKLEL 762
 Dnmt3a 812 QECLEHGRIAKFSKVRTITTRSNSIKQGDQHFVFMNEKEDILWCTEME 861
 |:||| | || ||.||||:||||:|||. | || | |::| ||||:|
 Dnmt3b 763 QDCLEFSRTAKLKKVQTITTKSNSIRQGNQLFPVVMNGKDDVLWCTELE 812
 Dnmt3a 862 RVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACV* 909
 |:|||| ||||| | ||:||||||| |||||:|||| |
 Dnmt3b 813 RIFGFPAHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE* 860

FIG.3A-2

DNMT3A 1 MPAMPSSGPGDTSSSSAAEREEDRKDGEEQEPRGKEERQEPSTTARKVGR
 DNMT3A 51 PGRKRKHPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPO
 DNMT3B 1MKGDRHLNGEEDAGGREDSILVNGACSDQSSDSP
 DNMT3A 101 PEEGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCTPKEGRGAPAEAG
 DNMT3B 36 PILEAIRTPAIRGGWASSRLSKREVSSLLSYTQDLTGDGDGEDGDGSDTP
 DNMT3A 151 KEQKETNIESMKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISK
 DNMT3B 86 VMPKLFRETRTRSESPAVRTRNNNSVSSRERHRPSRSTRGRQGRNHVDE
 DNMT3A 201 RKRDEWLARWKREAEKKAKVIAGMNAVEENQGPGESQKVEEASPPAVQQP
 DNMT3B 136 SPVEFPATRS�RRRATASAGTPWPSPSSYLITDLDDEDTH..GTPQS
 DNMT3A 251 TDPASPTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKL
 DNMT3B 184 SSTPYARLAQDSQQGMESPQVEADSGDGSSEYQDGKEFGIGDLVWGKI
 DNMT3A 301 RGFSWWPGRIVSWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSF
 DNMT3B 234 KGFSWWPAMVSWKATSKRQAMSGMRWVQWFGDGKFSEVSADKLVALGLF
 DNMT3A 351 CSAFHQATYNKQPMYRKAIYEV LQVASSRAGKLPVCHDSDES DTAKAVE
 DNMT3B 284 SQHFNLATFNKLVSYRKAMYLEKARVRAGKTFP.....SSPGDSLE
 DNMT3A 401 VQNKPMIEWALGGFQPSGPKGLEP....PEEKNPYKEVYTD MWVE....
 DNMT3B 327 DQLKPMLEWAHGGFKPTGIEGLKPNNTQPVVNKS KVRAGSRKLESRYKE
 DNMT3A 443PEAAAYAPPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRO
 DNMT3B 377 NKTRRRTADDSATSDYCPAPKRLKTNCYNNGKDRGDEDQSREQMASDVAN

FIG.3B-1

2

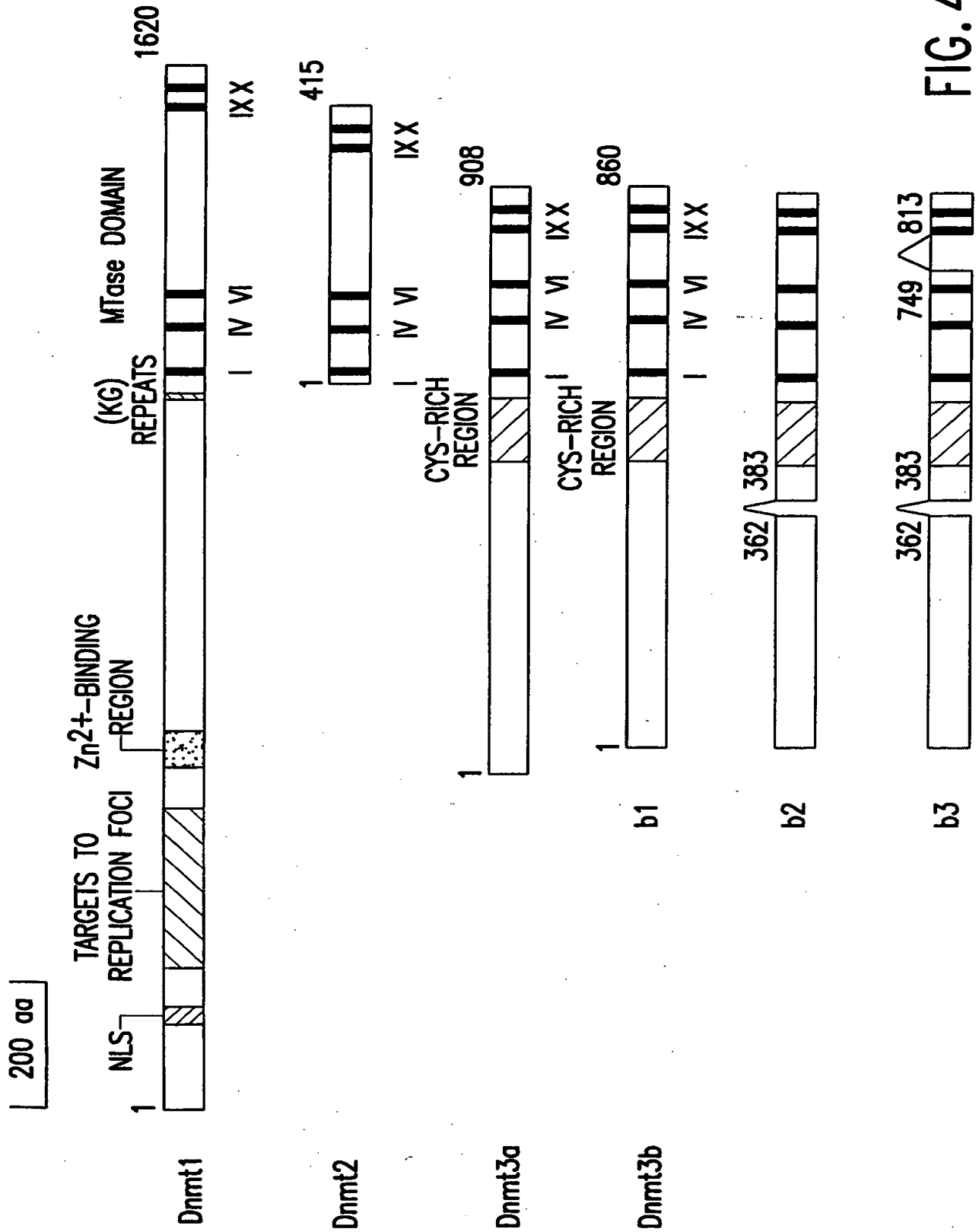


FIG. 4A

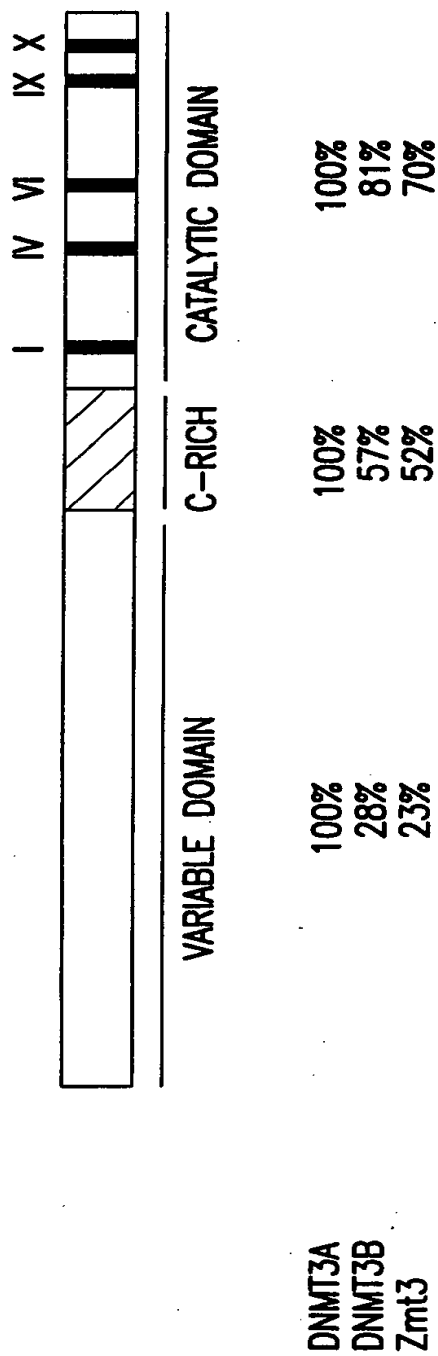
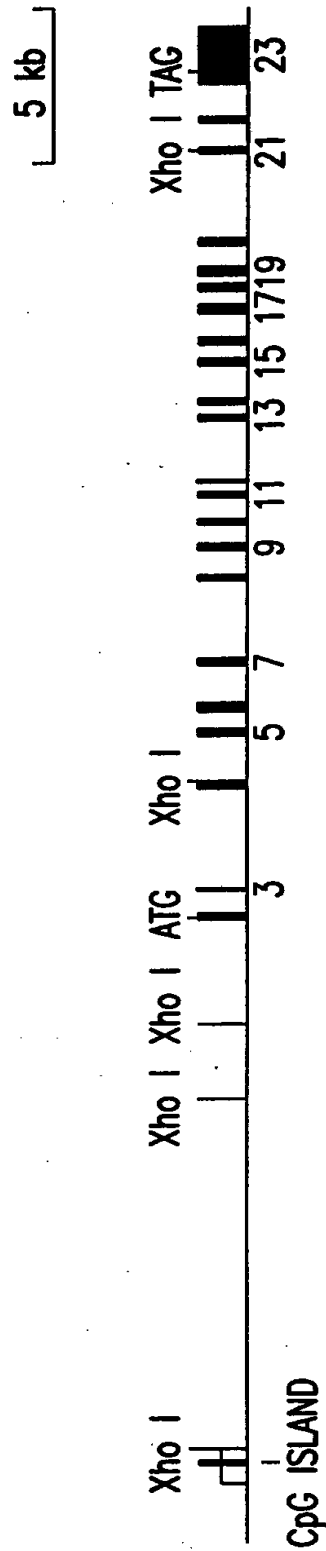


FIG. 4B



Exon1 (>=90bp) CGGCAGgtgagcgccccggggg.intron(17618bp). tggcttctccacagGAAAGC
 Exon2 (148bp) TCAGAGgtggctgggcagtgg.intron(887bp). CTGTTTCCTCTACAGGCCGAA
 Exon3 (62bp) ACACAGgtatggtctctgctc.intron(3343bp). tgttccctataaaagGACTTG
 Exon4 (102bp) CCAGCTgtaagtagccacacc.intron(1642bp). ctctcttgccttctagGTCCGA
 Exon5 (125bp) ACCAGGgttggtccccagatg.intron(602bp). tccttctgtccacagTCCCTG
 Exon6 (222bp) TATCAGgtatggccgagaggg.intron(1403bp). tgggttttcttccagGATGGG
 Exon7 (159bp) TCCGAGgtgagtcggggaag.intron(2588bp). gtcttctcttcttagGTCTCT
 Exon8 (108bp) CTGGAGgtaacatgggatgag.intron(917bp). actctgcctttgcagAAAGCT
 Exon9 (145bp) AACCAAGgtgggaatgagtcctc.intron(765bp). ttttccccctcaaaagTGGTTA
 Exon10 (60bp) AATACGgtatttccttccctgt.intron(1813bp). aattacctttcacagAGAACAA
 Exon11 (126bp) GCCGAGgtgattgttgggtac.intron(115bp). ttcttttctcaatagAACAAA
 Exon12 (45bp) TGGAAAGgtaacgttctctccc.intron(1095bp). ctgttttcttctacagATGGCT
 Exon13 (80bp) TGCCGGgttaagtctcctact.intron(417bp). ctctctggctgccagGATCGC
 Exon14 (113bp) CTGCCGgtgagcactgggccc.intron(1160bp). tgccactgggtccagGTGTTT
 Exon15 (184bp) GAATACgttaagccacaggtc.intron(600bp). ttcttacctggcagGAAGCC
 Exon16 (85bp) CGACAGgtgagttcggggaac.intron(824bp). ctctggccccccacagGCTACC
 Exon17 (146bp) AAAAATgtgagggcagctctgt.intron(536bp). gtctctctcttccagATTGAA
 Exon18 (91bp) TGTATGgtgagcatccttctc.intron(352bp). ctttctctgagcacagAGGGTA
 Exon19 (149bp) CTGGAGgtgaggaatctggg.intron(958bp). tcttctccccacagGTAAAT
 Exon20 (86bp) GAACAGgttaacaaagggtctc.intron(2867bp). ttgggtgttccccagGCCCGT
 Exon21 (70bp) GCCAAGgttaagaaagtacag.intron(801bp). cattttgttctccagTTAAAG
 Exon22 (119bp) CGAAAGgtgagcaaggctgca.intron(1434bp). ctccggtacccccagGATCTTT
 Exon23 (1585bp)

FIG.4D

I			IV			VI		
DnMT1	DVFSGCGGLSEGFHQAG	DVEMLCGGPPCQGFSGMNR	YRPRFFLL	ENVRN	FVSFKR			
Dnmt1	DVFSGCGGLSEGFHQAG	DVEMLCGGPPCQGFSGMNR	YRPRFFLL	ENVRN	FVSYRR			
MET1 (Ath)	DIFAGCGGLSHGLKKAG	QVDFINGGPPCQGFSGMNR	FRPRYFLL	ENVRT	FVSFNK			
Masc1	DTFCGGGVS LGARQAG	HVDILHLSPPCQTFSRAHT	VRPRLFTVE	ETDGI	MDRQS			
Masc2	DIFAGCGGLTLGLDLSG	EVDFIYGGPPCQGFSGMNR	YKPRFVLL	ENVKGL	ITTKL			
Dnmt2	EL YSGIGGMHHALRESH	SFNMILMSPPCQPFTRIGL	KLPKYILL	ENWKGE	FVSST			
M. Spr	SLFSGIGAFEAALRNIG	EFDLLVGGSPCQSFSVAGH	KQPKFFV	FENVKGL	INHDK			
DNMT3A	SLFDGIATGLLVKDLG	PFDLVIGGSPCNDLSIVNP	DRPFFWL	FENWAMG	VSDK			
Dnmt3a	SLFDGIATGLLVKDLG	PFDLVIGGSPCNDLSIVNP	DRPFFWL	FENWAMG	VSDK			
DNMT3B	SLFDGIATGYLVKELG	PFDLVIGGSPCNDLSNVNP	DRPFFWM	FENWAMK	VGDK			
Dnmt3b	SLFDGIATGYLVKELG	PFDLVIGGSPCNDLSNVNP	NRPFFWM	FENWAMK	VNDK			
Zmt3	SLFDGIATGYLVLRDLG	PFDLLIGGSPCNDLSIVNP	POPFFWL	FENVTFM	QTHVK			
consensus	--F-G-----G	-----GG-PC---S-N-	--P-F--ENV					

IX			X		
DNMT1	RWVSRECARSQGFP	LEGNILDKH	RQVGN	AVPP	LAKAIG
Dnmt1	RWVSRECARSQGFP	FFGNILDRH	RQVGN	AVPP	LAKAIG
MET1 (Ath)	RILTVRECARSQGFP	FAGNINH	KRIQI	GN	AVPP
Masc1	RKFTVRELACIQGFP	FVGTLDK	RRI	IGN	AVPP
Masc2	RVYTVRELARACGFP	GLGGVK	KWH	RNT	IGN
Dnmt2	RYFTPKEIANLQGFP	EKTTVK	QRYRLL	GN	SLNV
M. Spr	RRLTPLECFRLOAFD	AGISNS	QLYKQT	ENS	ITVT
DNMT3A	DILWCTEMERVFGFP	SNMSRL	ARQL	LLGR	SWSP
Dnmt3a	DILWCTEMERVFGFP	SNMSRL	ARQL	LLGR	SWSP
DNMT3B	DVLWCTELERIFGFP	SNMGRG	AROK	LLGR	SWSP
Dnmt3b	DVLWCTELERIFGFP	SNMGRG	AROK	LLGR	SWSP
Zmt3	DHIWITELEKIFGFP	KSMGRP	ORQR	V	LK
consensus	-----E-R--GFP	-----R-G-----P			

FIG. 5A

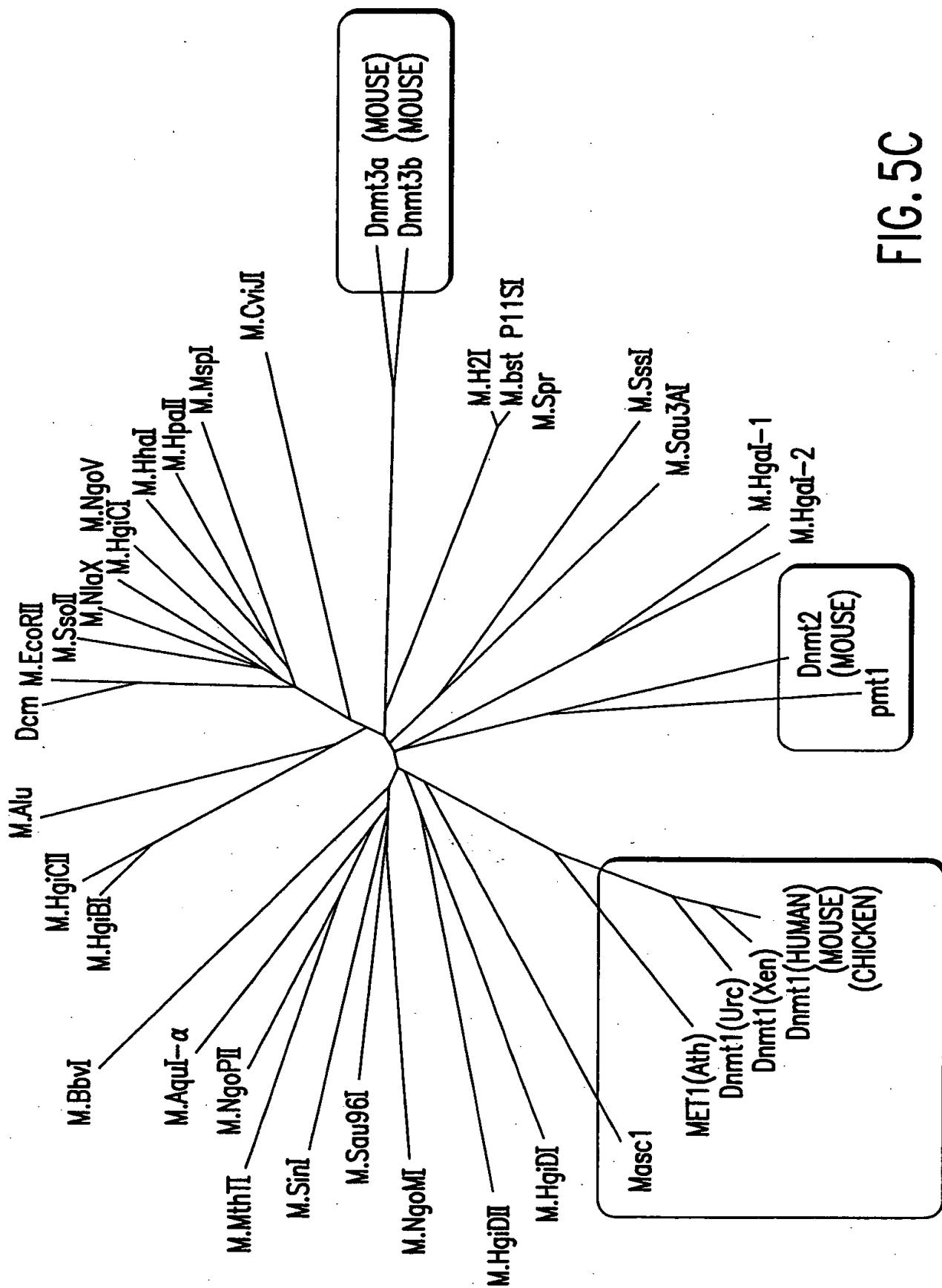


FIG. 5C

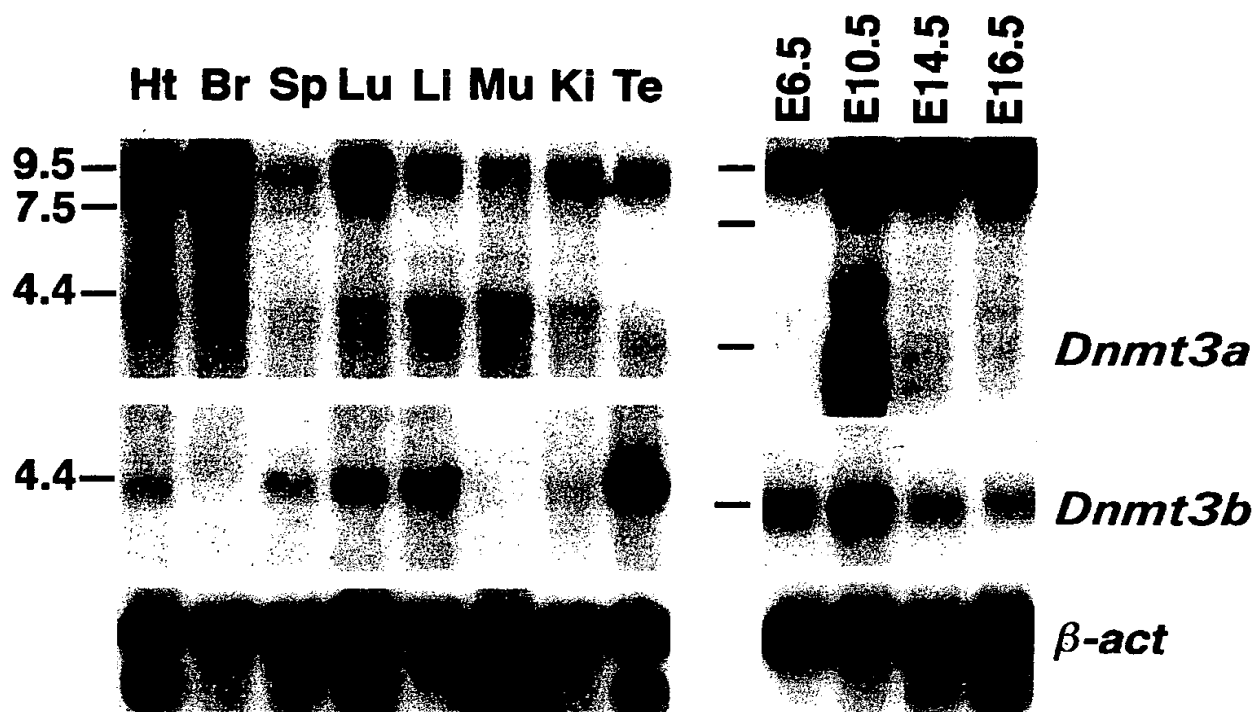


FIG.6A

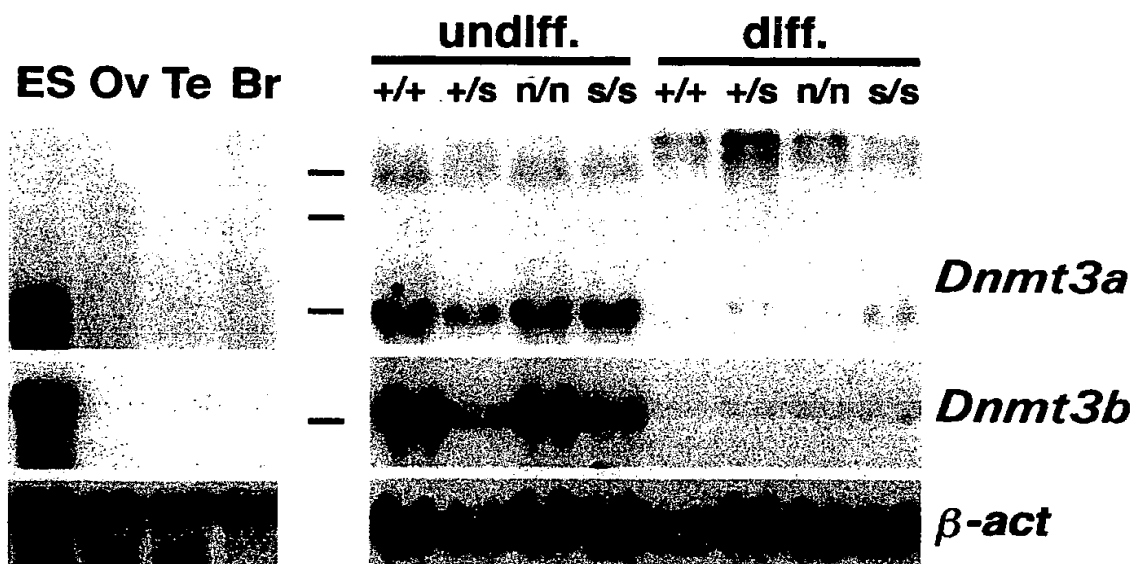


FIG.6B

FIG.6C

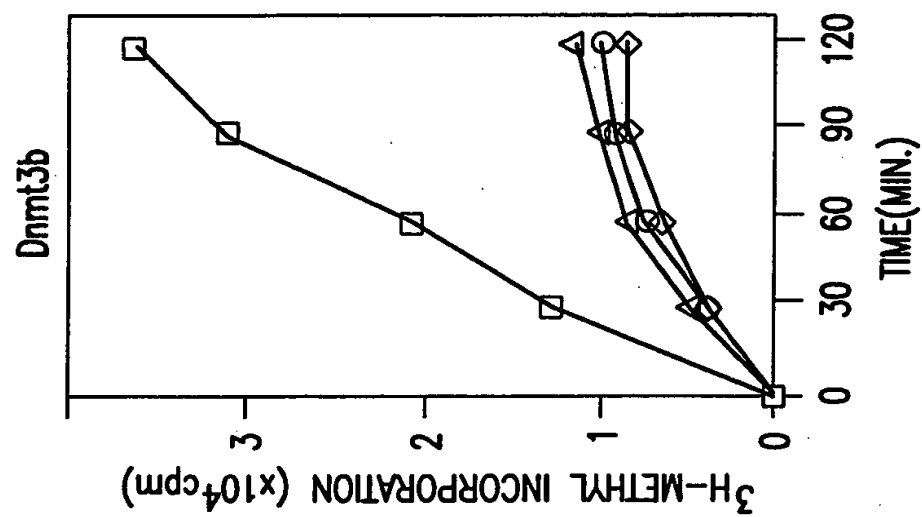


FIG. 7C

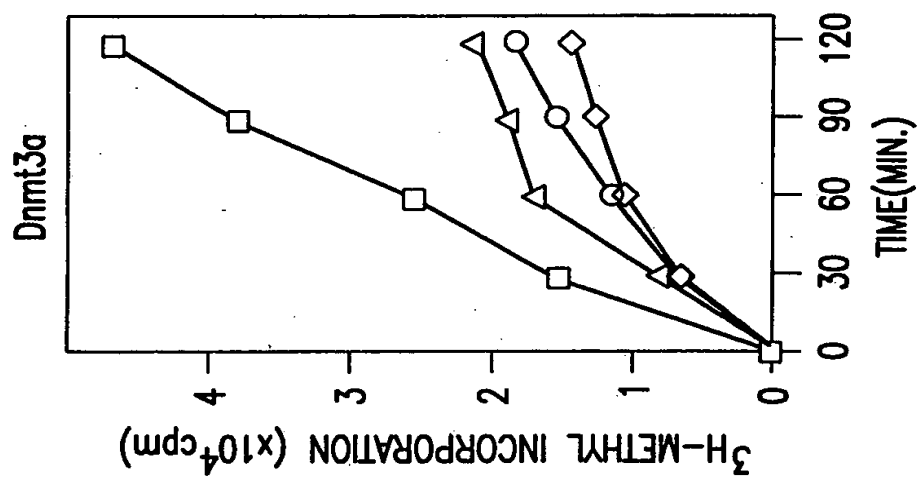


FIG. 7B

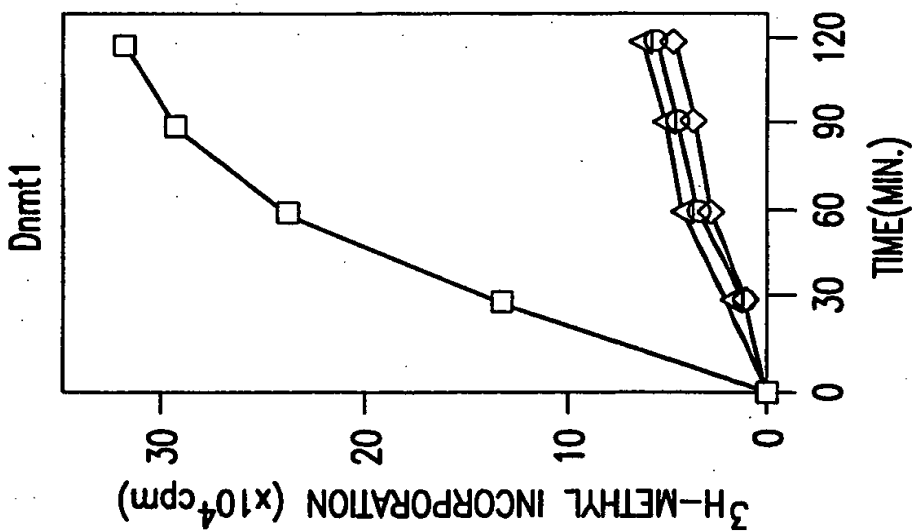


FIG. 7A



FIG. 7D

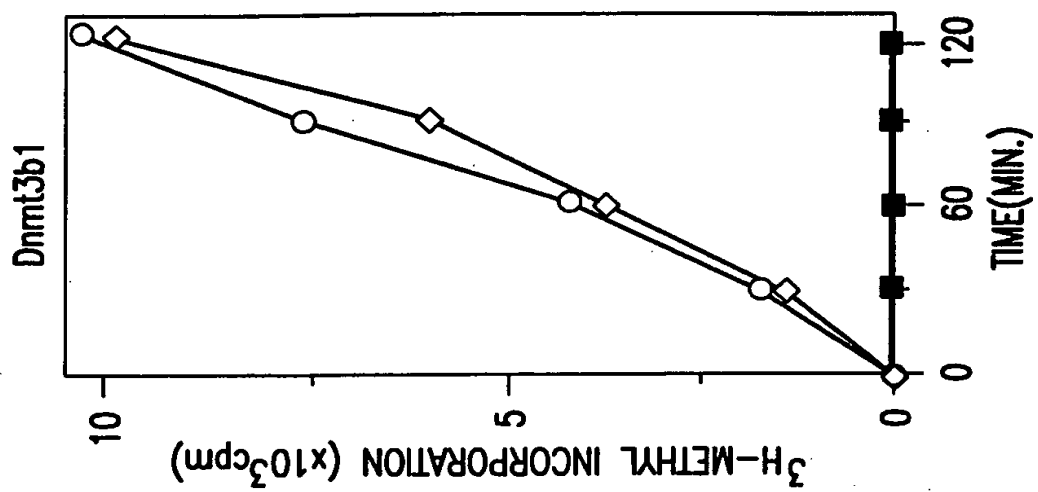


FIG. 8C

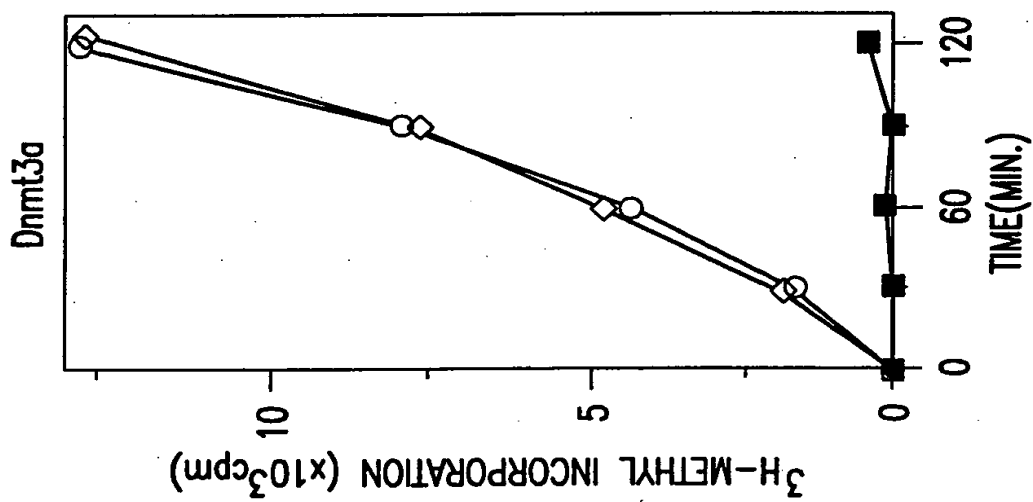


FIG. 8B

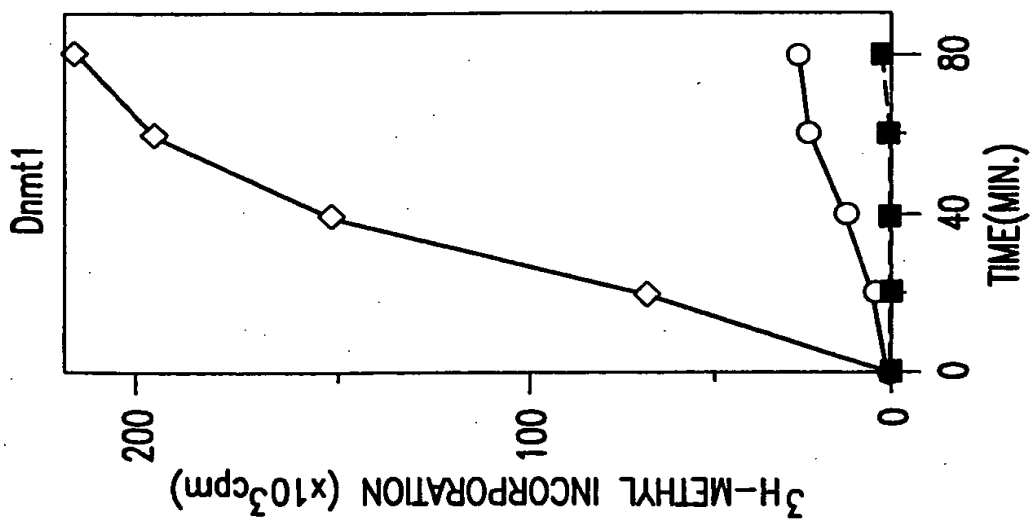


FIG. 8A

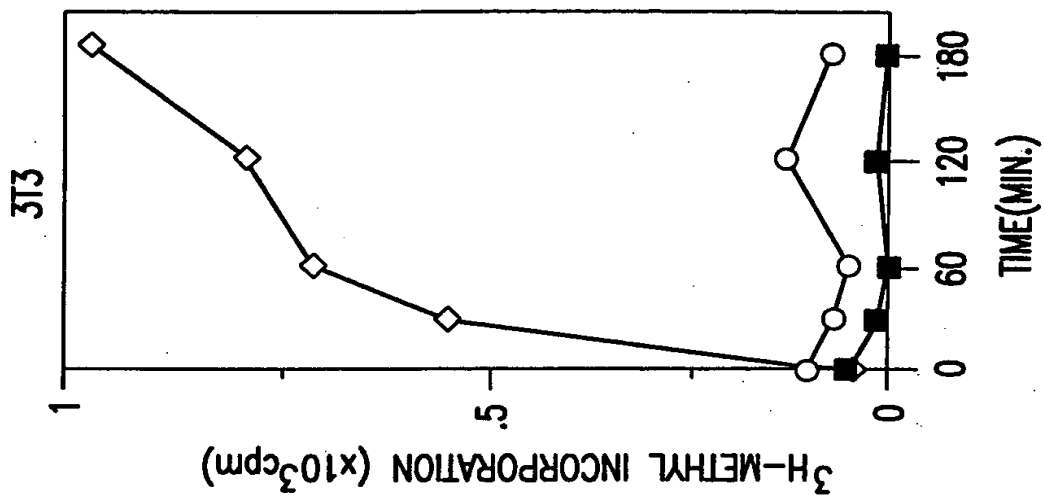


FIG. 8E

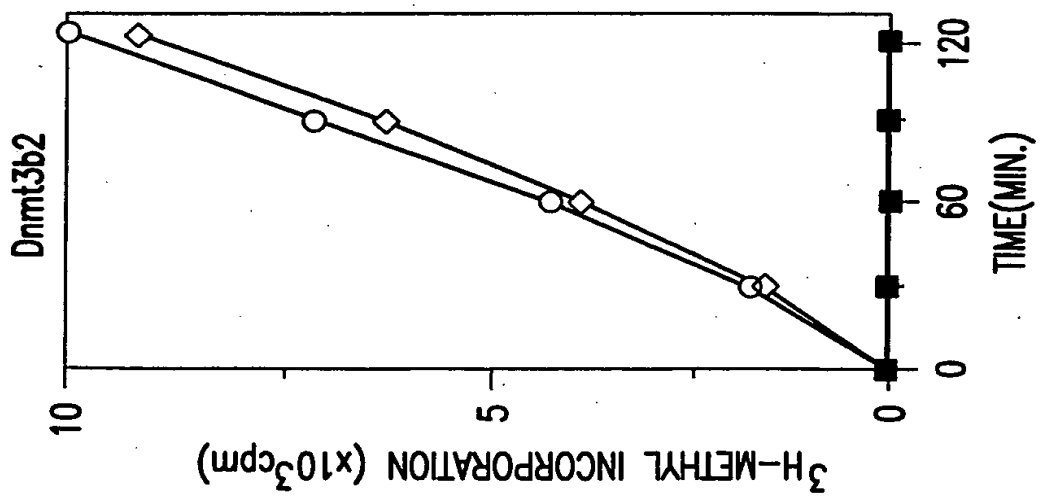
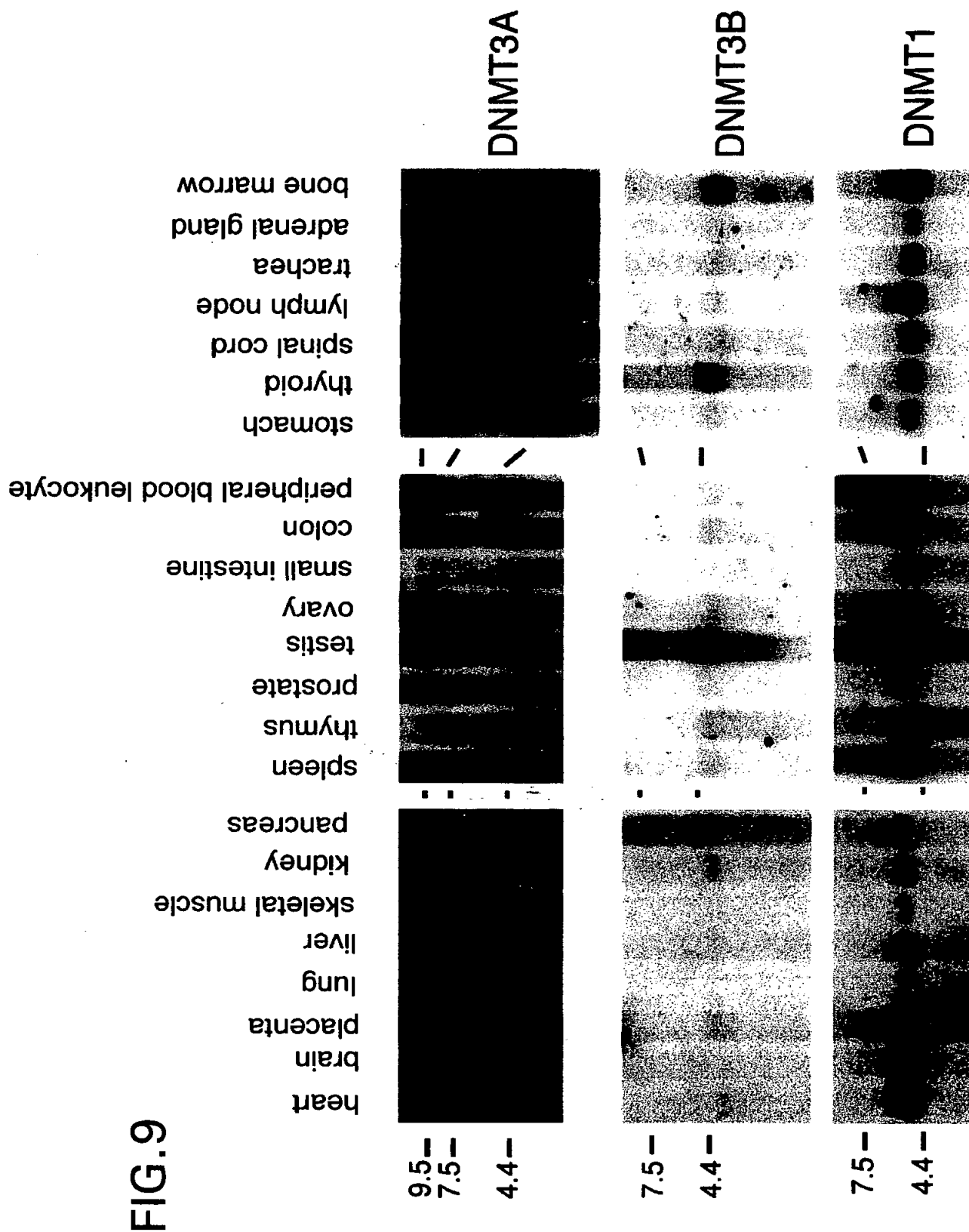


FIG. 8D



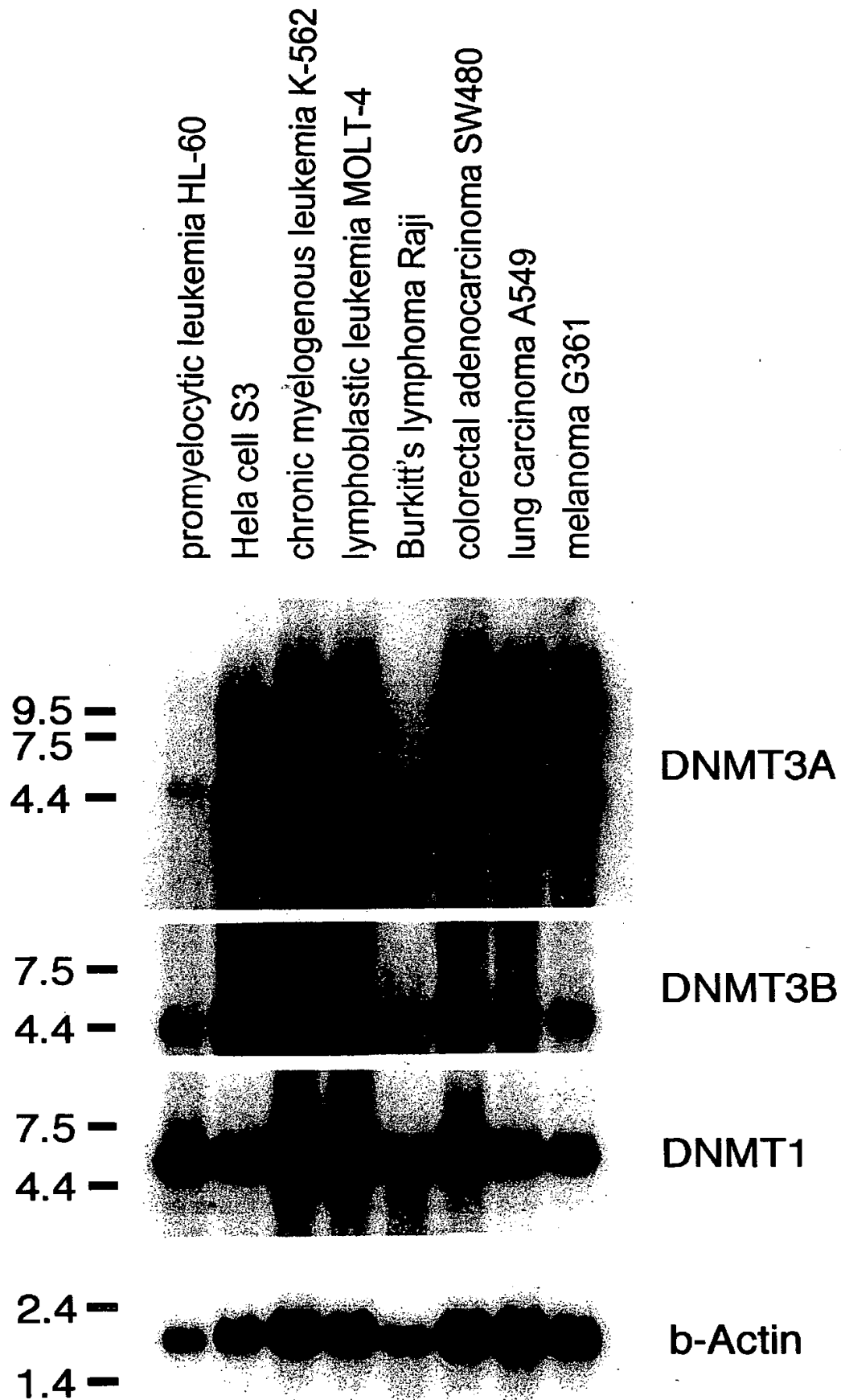


FIG. 10